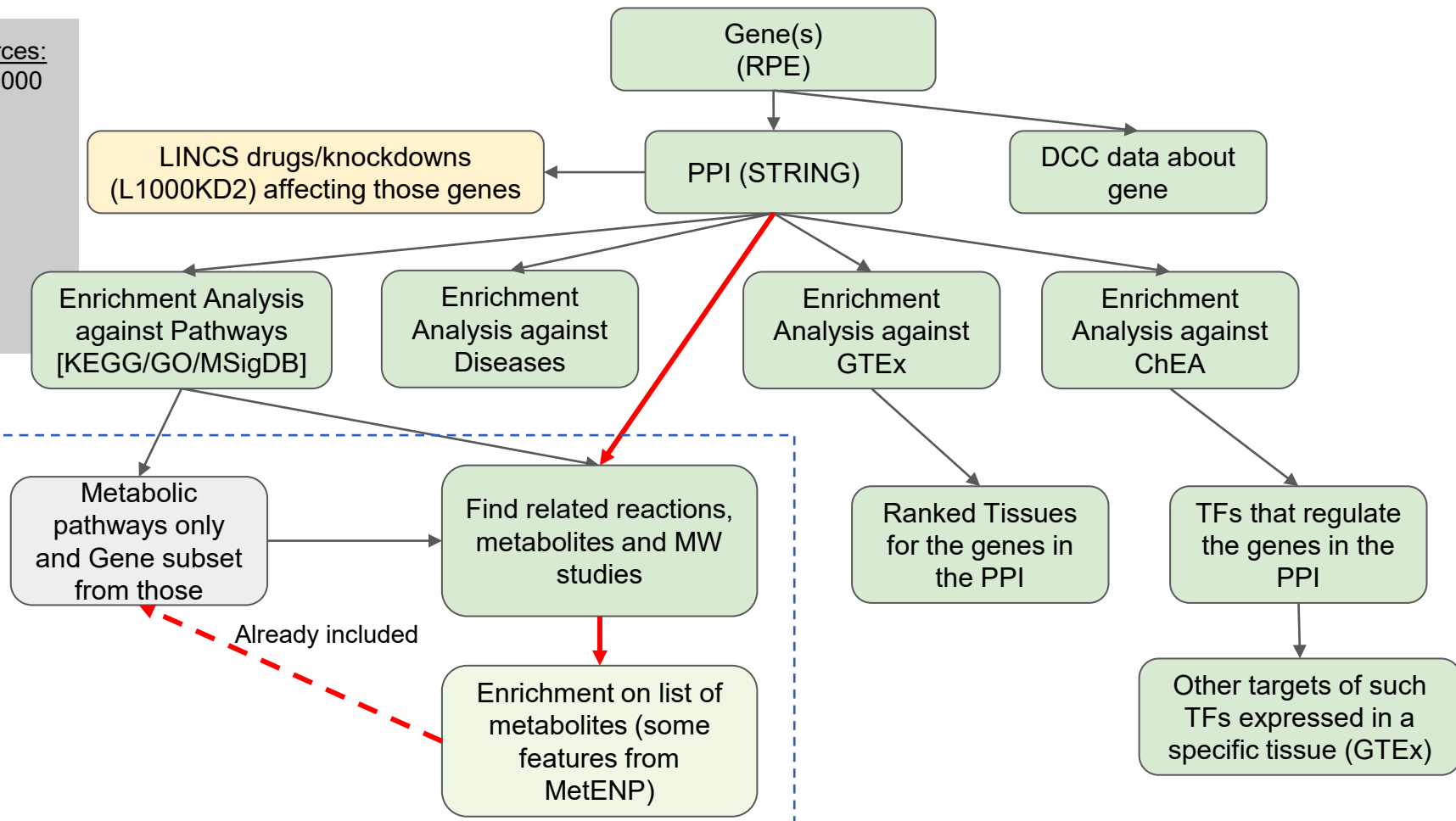


Development of Workflow using
Plyabook Workflow Builder Tool
for RPE as the seed gene
to Demonstrate use of **MetNet**

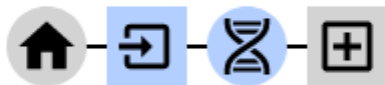
Use Case 11 – RPE related proteins/metabolites across DCCs (MW)

Data sources:

LINCS L1000
STRING
ChEA
GTEX
Enrichr
KEGG
GO
MSigDB
MW





Playbook Workflow Builder



Gene Input

The workflow starts with selecting a search term.

| | |
|---|--|
| Example  | Submit  |
|---|--|



Fetch StringDB PPI

The workflow starts with selecting RPE as the search term. For the given gene ID (SYMBOL), StringDB PPI was extracted using their API [1].

1. STRING api, <https://string-db.org/cgi/help.pl?subpage=api%23getting-all-the-string-interaction-partners-of-the-protein-set>

Species: hsa

Taxonomy ID: 9606

SYMBOL A SYMBOL B Score

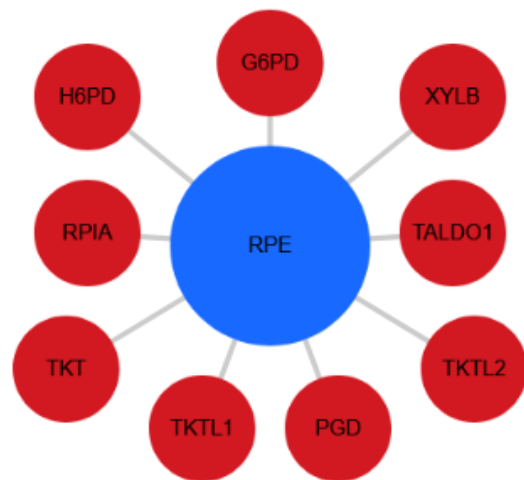
| | | |
|-----|--------|-------|
| RPE | RPIA | 0.998 |
| RPE | TKT | 0.997 |
| RPE | XYLB | 0.995 |
| RPE | PGD | 0.994 |
| RPE | TKTL2 | 0.991 |
| RPE | TKTL1 | 0.986 |
| RPE | TALDO1 | 0.981 |
| RPE | G6PD | 0.945 |
| RPE | H6PD | 0.943 |



Reformat StringDB PPI for plotting

The workflow starts with selecting RPE as the search term. For the given gene ID (SYMBOL), StringDB PPI was extracted using their API [1]. For the Given StringDB PPI, the list of nodes (GeneSet) is generated.

1. STRING api, <https://string-db.org/cgi/help.pl?subpage=api%23getting-all-the-string-interaction-partners-of-the-protein-set>





Given StringDB PPI, generate the list of nodes (GeneSet)

The workflow starts with selecting RPE as the search term. For the given gene ID (SYMBOL), StringDB PPI was extracted using their API [1]. For the Given StringDB PPI, the list of nodes (GeneSet) is generated.

1. STRING api, <https://string-db.org/cgi/help.pl?subpage=api%23getting-all-the-string-interaction-partners-of-the-protein-set>

| | Gene | | | |
|----|--------|--|--|--|
| 1 | RPE | | | |
| 2 | RPIA | | | |
| 3 | TKT | | | |
| 4 | XYLB | | | |
| 5 | PGD | | | |
| 6 | TKTL2 | | | |
| 7 | TKTL1 | | | |
| 8 | TALDO1 | | | |
| 9 | G6PD | | | |
| 10 | H6PD | | | |



Filter

Filter string...

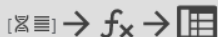
Output Type

- Signature [1 card]
- Gene [1 card]



Enrichr Enrichment Analysis

Perform Enrichment Analysis



Assemble GMT from Genesets

Group multiple independently generated gene sets into a single GMT



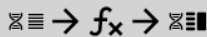
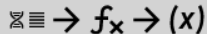
Filter genes by IDG Understudied Proteins

Based on IDG understudied proteins list



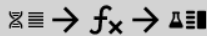
Filter genes by IDG GPCRs

Based on IDG understudied proteins list



Gene ID Conversion for a set of genes

Given one type of gene ID for a set of genes, generate other types of gene IDs.



MetGENE Reactions with GeneSet

Compute the MetGENE Reactions for a GeneSet



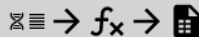
Filter genes by IDG Dark GPCRs

Based on IDG understudied proteins list



Filter genes by IDG Ion Channels

Based on IDG understudied proteins list



Perform CTD with Gene Set using KEGG

Get a CTD response for a set of genes using KEGG



MetGENE Studies with GeneSet

Compute the MetGENE studies function for a GeneSet



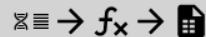
Filter genes by IDG Dark Ion Channels

Based on IDG understudied proteins list



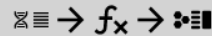
Filter genes by IDG Kinases

Based on IDG understudied proteins list



Perform CTD with Gene Set using STRING

Get a CTD response for a set of genes using STRING



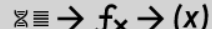
MetGENE Metabolites with GeneSet

Compute the MetGENE metabolites for a GeneSet



Filter genes by IDG Dark Kinases

Based on IDG understudied proteins list



Suggest a visualization method

This would visualize the Gene Set. Provide a description about what should be here.

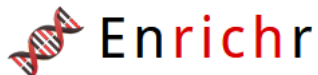


Enrichr Enrichment Analysis

The workflow starts with selecting RPE as the search term. For the given gene ID (SYMBOL), StringDB PPI was extracted using their API [1]. For the Given

StringDB PPI, the list of nodes (GeneSet) is generated.

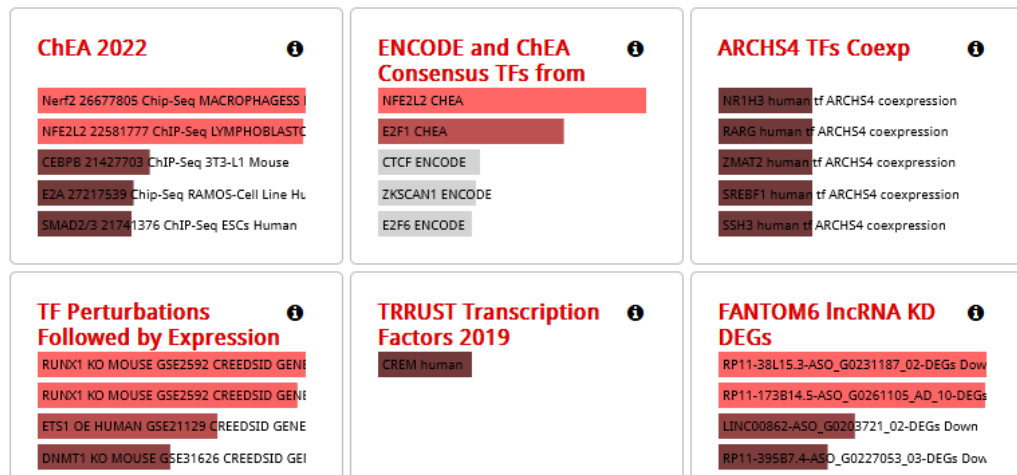
1. STRING api, <https://string-db.org/cgi/help.pl?subpage=api%23getting-all-the-string-interaction-partners-of-the-protein-set>

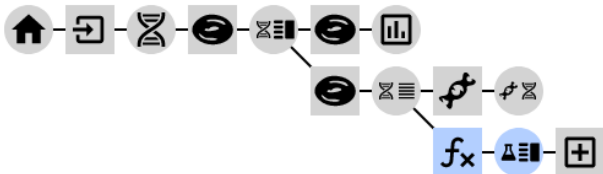


[Login](#) | [Register](#)

Transcription Pathways Ontologies Diseases/Drugs Cell Types Misc Legacy Crowd

Description **playbook-partnership (10 genes)**





MetGENE Reactions with GeneSet

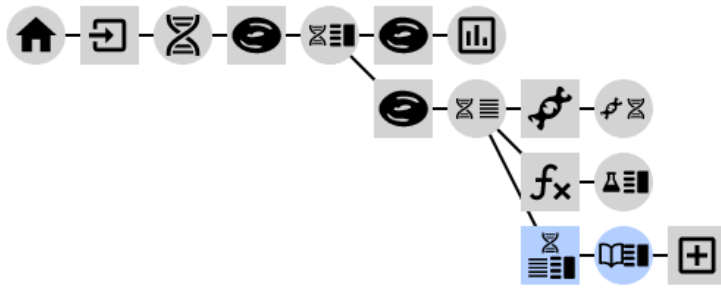
The workflow starts with selecting RPE as the search term. For the given gene ID (SYMBOL), StringDB PPI was extracted using their API [1]. For the Given StringDB PPI, the list of nodes (GeneSet) is generated. The geneset was then searched in the Metabolomics Workbench [2] to identify relevant reactions.

1. STRING api, <https://string-db.org/cgi/help.pl?subpage=api%23getting-all-the-string-interaction-partners-of-the-protein-set> 2. The Metabolomics Workbench, <https://www.metabolomicsworkbench.org/>

MetGENE Reactions

| Gene KEGG Rxn Id | KEGG Rxn Name | KEGG Rxn Equation |
|------------------|--|---|
| 6120 R01529 | D-Ribulose-5-phosphate 3-epimerase | D-Ribulose 5-phosphate <=> D-Xylulose 5-phosphate |
| Gene KEGG Rxn Id | KEGG Rxn Name | KEGG Rxn Equation |
| 22934 R01056 | D-ribose-5-phosphate aldose-ketose-isomerase | D-Ribose 5-phosphate <=> D-Ribulose 5-phosphate |

| Gene KEGG Rxn Id | KEGG Rxn Name | KEGG Rxn Equation |
|------------------|---|---|
| 7086 R01067 | D-Fructose 6-phosphate:D-glyceraldehyde-3-phosphate glycolaldehyde transferase | D-Fructose 6-phosphate + D-Glyceraldehyde 3-phosphate <=> D-Erythrose 4-phosphate + D-Xylulose 5-phosphate |
| 7086 R01641 | sedoheptulose-7-phosphate:D-glyceraldehyde-3-phosphate glycolaldehyde transferase | Sedoheptulose 7-phosphate + D-Glyceraldehyde 3-phosphate <=> D-Ribose 5-phosphate + D-Xylulose 5-phosphate |
| 7086 R01830 | beta-D-Fructose 6-phosphate:D-glyceraldehyde-3-phosphate glycolaldehyde transferase | beta-D-Fructose 6-phosphate + D-Glyceraldehyde 3-phosphate <=> D-Erythrose 4-phosphate + D-Xylulose 5-phosphate |
| 7086 R06590 | c(sedoheptulose-7-phosphate:D-glyceraldehyde-3-phosphate glycolaldehydetransferase; sedoheptulose-7-phosphate:D-glyceraldehyde-3-phosphate glycolaldehydetransferase) | Aminofructose 6-phosphate + D-Ribose 5-phosphate <=> Iminoerythrose 4-phosphate + Sedoheptulose 7-phosphate |
| Gene KEGG Rxn Id | KEGG Rxn Name | KEGG Rxn Equation |
| 9942 R01639 | ATP:D-xylulose 5-phosphotransferase | ATP + D-Xylulose <=> ADP + D-Xylulose 5-phosphate |



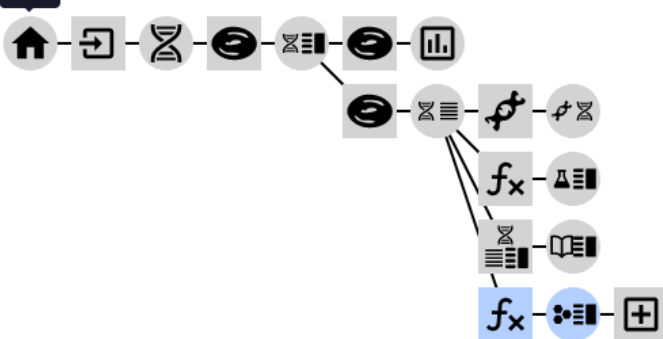
MetGENE Studies with GeneSet

The workflow starts with selecting RPE as the search term. For the given gene ID (SYMBOL), StringDB PPI was extracted using their API [1]. For the Given StringDB PPI, the list of nodes (GeneSet) is generated. The geneset was then searched in the Metabolomics Workbench [2] to identify relevant studies related to the genes.

1. STRING api, <https://string-db.org/cgi/help.pl?subpage=api%23getting-all-the-string-interaction-partners-of-the-protein-set> 2. The Metabolomics Workbench, <https://www.metabolomicsworkbench.org/>

MetGENE Studies

| Compound Id | Refmet Name | Study Ids |
|-------------|----------------------|---|
| C00199 | Ribulose 5-phosphate | ST002719, ST002584, ST002555, ST002505, ST001983, ST001842, ST001527, ST001167, ST001136, ST001135, ST001000230, ST0000087, ST0000058 |
| C00231 | Xylulose 5-phosphate | ST002584, ST002328, ST002234, ST001167, ST001136, ST001135, ST000231, ST000230, ST002805, ST002702, ST002584, ST002567, ST002536, ST002505, ST002323, ST002246, ST002234, ST002115, ST001788, ST001780, ST001720, ST001692, ST001441, ST001420, ST001402, ST001398, ST001384, ST001335, ST001135, ST001133, ST001074, ST001039, ST001004, ST000923, ST000876, ST000842, ST000841, ST000839, ST000719, ST000714, ST000600, ST000617, ST000520, ST000450, ST000448, ST000404, ST000403, ST000370, ST000 |
| C00117 | Ribose 5-phosphate | |



MetGENE Metabolites with GeneSet

The workflow starts with selecting RPE as the search term. For the given gene ID (SYMBOL), StringDB PPI was extracted using their API [1]. For the Given StringDB PPI, the list of nodes (GeneSet) is generated. The geneset was then searched in the Metabolomics Workbench [Metabolomics Workbench, 2] to identify associated metabolites.

1. STRING api, <https://string-db.org/cgi/help.pl?subpage=api%23getting-all-the-string-interaction-partners-of-the-protein-set> 2. <https://www.metabolomicsworkbench.org/>

MetGENE metabolites

[Full table as a json file is available at MetGENE metabolites json file. The tables for the first 9 genes are shown below.](#)

| Gene | KEGG_COMPOUND_ID | REFMET_NAME | KEGG_REACTION_ID | METSTAT_LINK |
|------|------------------|----------------------|------------------|--------------|
| | 6120 C00199 | Ribulose 5-phosphate | R01529 | METSTAT |
| | 6120 C00231 | Xylulose 5-phosphate | R01529 | METSTAT |
| Gene | KEGG_COMPOUND_ID | REFMET_NAME | KEGG_REACTION_ID | METSTAT_LINK |
| | 22934 C00117 | Ribose 5-phosphate | R01056 | METSTAT |
| | 22934 C00199 | Ribulose 5-phosphate | R01056 | METSTAT |
| Gene | KEGG_COMPOUND_ID | REFMET_NAME | KEGG_REACTION_ID | METSTAT_LINK |
| | 7086 C00085 | Fructose 6-phosphate | R01067 | METSTAT |

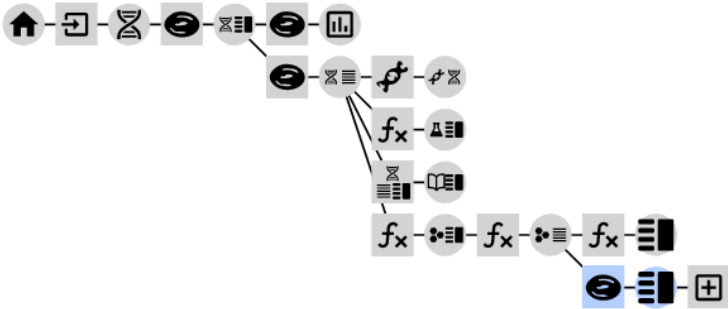
Extract metabolite set information

The workflow starts with selecting RPE as the search term. For the given gene ID (SYMBOL), StringDB PPI was extracted using their API [1]. For the Given StringDB PPI, the list of nodes (GeneSet) is generated. The geneset was then searched in the Metabolomics Workbench [Metabolomics Workbench, 2] to identify associated metabolites. Then, MetaboliteSet (REFMET names) is extracted from the table of MetGENE Metabolites for various genes. The metabolites were then searched in the Metabolomics Workbench [3] to extract more information about the metabolites.

1. STRING api, <https://string-db.org/cgi/help.pl?subpage=api%23getting-all-the-string-interaction-partners-of-the-protein-set> 2.
- <https://www.metabolomicsworkbench.org/> 3. The Metabolomics Workbench,
- <https://www.metabolomicsworkbench.org/>

Metabolite Summary

| Name | RefMet Name | PubChem CID | Exact Mass | Formula | Super class | Main class | Sub class | InChI |
|------|------------------------------|-------------|------------|---------------|---------------|----------------------------|-------------------|-------|
| | Ribulose 5-phosphate | | 230.0192 | C5H11O8P | Carbohydrates | Monosaccharide phosphates | Monosaccharides | |
| | Xylulose 5-phosphate | | 230.0192 | C5H11O8P | Carbohydrates | Monosaccharide phosphates | Monosaccharides | |
| | Ribose 5-phosphate | | 230.0192 | C5H11O8P | Carbohydrates | Monosaccharide phosphates | Monosaccharides | |
| | Fructose 6-phosphate | | 260.0297 | C6H13O9P | Carbohydrates | Monosaccharide phosphates | Monosaccharides | |
| | Glyceraldehyde 3-phosphate | | 169.9980 | C3H7O6P | Organic acids | Short-chain acids | Short-chain acids | |
| | Erythrose 4-phosphate | | 200.0086 | C4H9O7P | Carbohydrates | Monosaccharide phosphates | Monosaccharides | |
| | Sedoheptulose 7-phosphate | | 290.0403 | C7H15O10P | Carbohydrates | Monosaccharide phosphates | Monosaccharides | |
| | ATP | | 506.9958 | C10H16N5O13P3 | Nucleic acids | Purine rNTP | Purines | |
| | Adenosine | | 267.0968 | C10H13N5O4 | Nucleic acids | Purine ribonucleosides | Purines | |
| | Xylulose | | 150.0528 | C5H10O5 | Carbohydrates | Pentoses | Monosaccharides | |
| | NAD+ | | 663.1091 | C21H27N7O14P2 | Nucleic acids | Nicotinamide dinucleotides | Nicotinamides | |
| | NADH | | 665.1248 | C21H29N7O14P2 | Nucleic acids | Nicotinamide dinucleotides | Nicotinamides | |
| | NADPH | | 745.0911 | C21H30N7O17P3 | Nucleic acids | Nicotinamide dinucleotides | Nicotinamides | |
| | NADP+ | | 743.0755 | C21H28N7O17P3 | Nucleic acids | Nicotinamide dinucleotides | Nicotinamides | |
| | 6-Phosphogluconic acid | | 276.0246 | C6H13O10P | Carbohydrates | Monosaccharide phosphates | Monosaccharides | |
| | Glucose 6-phosphate | | 260.0297 | C6H13O9P | Carbohydrates | Monosaccharide phosphates | Monosaccharides | |
| | 6-Phosphonoglucono-D-lactone | | 258.0141 | C6H11O9P | Carbohydrates | Monosaccharide phosphates | Monosaccharides | |



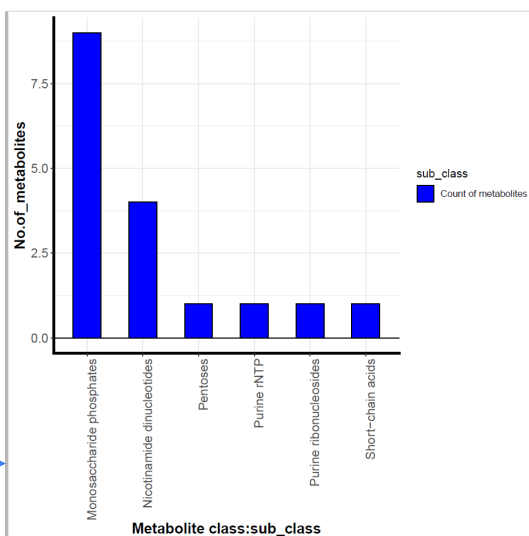
Perform MetENP analysis on a list of metabolites (no numeric data)

The workflow starts with selecting RPE as the search term. For the given gene ID (SYMBOL), STRINGDB PPI was extracted using their API [1]. For the Given STRINGDB PPI, the list of nodes (GeneSet) is generated. The geneset was then searched in the Metabolomics Workbench [Metabolomics Workbench, 2] to identify associated metabolites. Then, MetaboliteSet (REFMET names) is extracted from the table of MetGENE Metabolites for various genes. For the given list of metabolites, MetENP analysis is carried out and the result (a list of output files) is displayed as a table for further exploration.

1. STRING api, <https://string-db.org/cgi/help.pl?subpage=api%23getting-all-the-string-interaction-partners-of-the-protein-set.2>.
<https://www.metabolomicsworkbench.org/>

- | No. | Description/URL |
|-----|--|
| 0 | Plot of metabolite count |
| 1 | Metabolite class enrichment plot |
| 2 | Metabolite class enrichment table |
| 3 | Metabolite -> Reaction -> Pathway information |
| 4 | KEGG pathway enrichment |
| 5 | Metabolite KEGG pathway network |
| 6 | Metabolite pathway heatmap |
| 7 | Metabolite pathway dot plot |
| 8 | Orthology centric information on Metabolite related genes |
| 9 | Metabolite related gene names and Entrez IDs |
| 10 | Reaction-Substrate information for the Metabolites/related genes |
| 11 | PPI for the metabolite related genes: txt |
| 12 | PPI for the metabolite related genes: json |

In the future, these component results will be separated so that additional transformations can be applied.



```

https://bdcw.org//MetENP/out/673ca_20230914_113316/met_gene_only.json

JSON Raw Data Headers
Save Copy Collapse All Expand All Filter JSON

0:
  gene_name: "aldo-keto reductase family 1 member A1"
  gene_id: "10327"
1:
  gene_name: "glycerol-3-phosphate dehydrogenase 1 like"
  gene_id: "23171"
2:
  gene_name: "glycerol-3-phosphate dehydrogenase 1"
  gene_id: "2819"
3:
  gene_name: "sorbitol dehydrogenase"
  gene_id: "6652"
4:
  gene_name: "aldo-keto reductase family 1 member B"
  gene_id: "231"
5:
  gene_name: "aldo-keto reductase family 1 member B15"
  gene_id: "441282"
6:
  gene_name: "aldo-keto reductase family 1 member B18"

```

